Discussion of "An extension of the unified skew-normal family of distributions and applications to Bayesian binary regression"

by Brunero Liseo (joint work with Paolo Onorati)

Athanasios Kottas

Department of Statistics, University of California, Santa Cruz

OBayes 2022, UC Santa Cruz, September 6-10, 2022

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \boldsymbol{\beta} \overset{ind.}{\sim} \text{Bernoulli}(\Lambda(\boldsymbol{x}_i^T \boldsymbol{\beta}))$, $i = 1, \ldots, n$.
  - ▷ The pSUN is a conjugate prior for $\boldsymbol{\beta}$, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \boldsymbol{\beta} \stackrel{ind.}{\sim} \text{Bernoulli}(\Lambda(\boldsymbol{x}_i^T\boldsymbol{\beta})), i = 1, \ldots, n.$
  - ▷ The pSUN is a conjugate prior for $\boldsymbol{\beta}$, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \beta \overset{ind.}{\sim} \text{Bernoulli}(\Lambda(x_i^T\beta)), i = 1, \ldots, n$.
  - ▷ The pSUN is a conjugate prior for β, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \beta \overset{ind.}{\sim} \text{Bernoulli}(\Lambda(x_i^T\beta)), i = 1, \ldots, n$.
  - ▷ The pSUN is a conjugate prior for β, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \beta \overset{ind.}{\sim} \text{Bernoulli}(\Lambda(x_i^T \beta))$, $i = 1, \ldots, n$.
  - ▷ The pSUN is a conjugate prior for β, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \boldsymbol{\beta} \overset{ind.}{\sim} \text{Bernoulli}(\Lambda(\boldsymbol{x}_i^T\boldsymbol{\beta}))$, $i = 1, \ldots, n$.
  - ▷ The pSUN is a conjugate prior for $\boldsymbol{\beta}$, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Summary

- Extension of the unified skew-normal (SUN) distribution (Arellano-Valle & Azzalini, 2006).
- Perturbed unified skew-normal (pSUN) distribution.
  - ▷ Replace the mulitvariate normal variables that drive a stochastic representation of the SUN distribution with scale mixtures of normals.
  - ▷ Gibbs sampler to simulate from the pSUN distribution.
  - ▷ Motivation: explore a general class of conjugate priors for the regression coefficients in binary regression models.
  - ▷ A bit difficult to envision incorporating prior beliefs into the general version of the pSUN distribution (or the SUN distribution for that matter), but useful priors are included as special cases.
- Binary regression model: $y_i \mid \boldsymbol{\beta} \overset{ind.}{\sim} \text{Bernoulli}(\Lambda(\boldsymbol{x}_i^T \boldsymbol{\beta})), i = 1, \ldots, n$.
  - ▷ The pSUN is a conjugate prior for $\boldsymbol{\beta}$, provided the inverse link Λ is the c.d.f. of a scale normal mixture.

Bayesian binary regression

- MCMC with data augmentation (normal prior for β)
  - Probit link: $\beta \mid z, \text{data} \sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\beta \mid z, u, \text{data} \sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \overset{ind.}{\sim} \mathrm{N}(x_i^T\beta, (2u_i)^2)$, and $u_i \overset{i.i.d.}{\sim} \mathrm{KS}$ (Holmes & Held, 2006).
  - Logit link: $\beta \mid v, \text{data} \sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).
- Under the probit link, and a normal prior, $\beta \mid \text{data} \sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - ▷ More generally, conjugate SUN prior for β under the probit link.
  - ▷ Independent sampling from $p(\beta \mid \text{data})$ (practical for small/moderate n).
- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - ▷ Extends the story to symmetric links, including the probit and logit.
  - ▷ Requires a Gibbs sampler to explore $p(\beta \mid \text{data})$.

Bayesian binary regression

- MCMC with data augmentation (normal prior for $\boldsymbol{\beta}$)
  - Probit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \text{data} \sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \boldsymbol{u}, \text{data} \sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \overset{ind.}{\sim} \mathrm{N}(\boldsymbol{x}_i^T\boldsymbol{\beta}, (2u_i)^2)$, and $u_i \overset{i.i.d.}{\sim}$ KS (Holmes & Held, 2006).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{v}, \text{data} \sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).
- Under the probit link, and a normal prior, $\boldsymbol{\beta} \mid \text{data} \sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - ▷ More generally, conjugate SUN prior for $\boldsymbol{\beta}$ under the probit link.
  - ▷ Independent sampling from $p(\boldsymbol{\beta} \mid \text{data})$ (practical for small/moderate n).
- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - ▷ Extends the story to symmetric links, including the probit and logit.
  - ▷ Requires a Gibbs sampler to explore $p(\boldsymbol{\beta} \mid \text{data})$.

Bayesian binary regression

- MCMC with data augmentation (normal prior for $\boldsymbol{\beta}$)
  - Probit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \text{data} \sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \boldsymbol{u}, \text{data} \sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \overset{ind.}{\sim} \mathrm{N}(\boldsymbol{x}_i^T\boldsymbol{\beta}, (2u_i)^2)$, and $u_i \overset{i.i.d.}{\sim} \mathrm{KS}$ (Holmes & Held, 2006).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{v}, \text{data} \sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).
- Under the probit link, and a normal prior, $\boldsymbol{\beta} \mid \text{data} \sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - ▷ More generally, conjugate SUN prior for $\boldsymbol{\beta}$ under the probit link.
  - ▷ Independent sampling from $p(\boldsymbol{\beta} \mid \text{data})$ (practical for small/moderate n).
- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - ▷ Extends the story to symmetric links, including the probit and logit.
  - ▷ Requires a Gibbs sampler to explore $p(\boldsymbol{\beta} \mid \text{data})$.

Bayesian binary regression

- MCMC with data augmentation (normal prior for $\boldsymbol{\beta}$)
  - Probit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \text{data} \sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \boldsymbol{u}, \text{data} \sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \overset{ind.}{\sim} \mathrm{N}(\boldsymbol{x}_i^T \boldsymbol{\beta}, (2u_i)^2)$, and $u_i \overset{i.i.d.}{\sim}$ KS (Holmes & Held, 2006).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{v}, \text{data} \sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).

- Under the probit link, and a normal prior, $\boldsymbol{\beta} \mid \text{data} \sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - More generally, conjugate SUN prior for $\boldsymbol{\beta}$ under the probit link.
  - Independent sampling from $p(\boldsymbol{\beta} \mid \text{data})$ (practical for small/moderate n).

- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - Extends the story to symmetric links, including the probit and logit.
  - Requires a Gibbs sampler to explore $p(\boldsymbol{\beta} \mid \text{data})$.

Bayesian binary regression

- MCMC with data augmentation (normal prior for $\boldsymbol{\beta}$)
  - Probit link: $\boldsymbol{\beta} \mid \boldsymbol{z}$, data $\sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \boldsymbol{u}$, data $\sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \overset{ind.}{\sim} \mathrm{N}(\boldsymbol{x}_i^T\boldsymbol{\beta}, (2u_i)^2)$, and $u_i \overset{i.i.d.}{\sim}$ KS (Holmes & Held, 2006).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{v}$, data $\sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).

- Under the probit link, and a normal prior, $\boldsymbol{\beta} \mid$ data $\sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - ▷ More generally, conjugate SUN prior for $\boldsymbol{\beta}$ under the probit link.
  - ▷ Independent sampling from $p(\boldsymbol{\beta} \mid \text{data})$ (practical for small/moderate n).

- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - ▷ Extends the story to symmetric links, including the probit and logit.
  - ▷ Requires a Gibbs sampler to explore $p(\boldsymbol{\beta} \mid \text{data})$.

Bayesian binary regression

- MCMC with data augmentation (normal prior for $\boldsymbol{\beta}$)
  - Probit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \text{data} \sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \boldsymbol{u}, \text{data} \sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \stackrel{ind.}{\sim} \mathrm{N}(\boldsymbol{x}_i^T\boldsymbol{\beta}, (2u_i)^2)$, and $u_i \stackrel{i.i.d.}{\sim} \mathrm{KS}$ (Holmes & Held, 2006).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{v}, \text{data} \sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).

- Under the probit link, and a normal prior, $\boldsymbol{\beta} \mid \text{data} \sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - ▷ More generally, conjugate SUN prior for $\boldsymbol{\beta}$ under the probit link.
  - ▷ Independent sampling from $p(\boldsymbol{\beta} \mid \text{data})$ (practical for small/moderate n).

- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - ▷ Extends the story to symmetric links, including the probit and logit.
  - ▷ Requires a Gibbs sampler to explore $p(\boldsymbol{\beta} \mid \text{data})$.

Bayesian binary regression

- MCMC with data augmentation (normal prior for $\boldsymbol{\beta}$)
  - Probit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \text{data} \sim \mathrm{N}_p$, given normally distributed latent variables z_i (Albert & Chib, 1993).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{z}, \boldsymbol{u}, \text{data} \sim \mathrm{N}_p$, given latent variables $z_i \mid u_i \overset{ind.}{\sim} \mathrm{N}(\boldsymbol{x}_i^T \boldsymbol{\beta}, (2u_i)^2)$, and $u_i \overset{i.i.d.}{\sim}$ KS (Holmes & Held, 2006).
  - Logit link: $\boldsymbol{\beta} \mid \boldsymbol{v}, \text{data} \sim \mathrm{N}_p$, given PG latent variables v_i (Polson et al., 2013).

- Under the probit link, and a normal prior, $\boldsymbol{\beta} \mid \text{data} \sim \mathrm{SUN}_{p,n}$ (Durante, 2019).
  - ▷ More generally, conjugate SUN prior for $\boldsymbol{\beta}$ under the probit link.
  - ▷ Independent sampling from $p(\boldsymbol{\beta} \mid \text{data})$ (practical for small/moderate n).

- New contribution: results for the pSUN prior (Onorati & Liseo, 2022).
  - ▷ Extends the story to symmetric links, including the probit and logit.
  - ▷ Requires a Gibbs sampler to explore $p(\boldsymbol{\beta} \mid \text{data})$.

A few comments

- Expressions for $E(\beta \mid \text{data})$ or $E(\beta_k \mid \text{data})$ under special cases of the prior? Can they be efficiently computed without pSUN sampling?

- Computing: general approach or algorithms tailored to important special cases? software?

- Extensions?
  - Semiparametric model: nonparametric scale normal mixture for the inverse link + pSUN prior for β.
  - General model settings with binary regression as a component (binary longitudinal responses, nonparametric mixtures, ordinal regression ...)

- Many thanks to Brunero for his contribution to the conference, and congrats to Paolo & Brunero for a nice paper!

A few comments

- Expressions for $\mathrm{E}(\boldsymbol{\beta} \mid \text{data})$ or $\mathrm{E}(\beta_k \mid \text{data})$ under special cases of the prior? Can they be efficiently computed without pSUN sampling?

- Computing: general approach or algorithms tailored to important special cases? software?

- Extensions?
  - Semiparametric model: nonparametric scale normal mixture for the inverse link + pSUN prior for $\boldsymbol{\beta}$.
  - General model settings with binary regression as a component (binary longitudinal responses, nonparametric mixtures, ordinal regression ...)

- Many thanks to Brunero for his contribution to the conference, and congrats to Paolo & Brunero for a nice paper!

A few comments

- Expressions for $E(\boldsymbol{\beta} \mid \text{data})$ or $E(\beta_k \mid \text{data})$ under special cases of the prior? Can they be efficiently computed without pSUN sampling?

- Computing: general approach or algorithms tailored to important special cases? software?

- Extensions?

  - Semiparametric model: nonparametric scale normal mixture for the inverse link + pSUN prior for $\boldsymbol{\beta}$.

  - General model settings with binary regression as a component (binary longitudinal responses, nonparametric mixtures, ordinal regression ...)

- Many thanks to Brunero for his contribution to the conference, and congrats to Paolo & Brunero for a nice paper!

A few comments

- Expressions for $\mathrm{E}(\boldsymbol{\beta} \mid \text{data})$ or $\mathrm{E}(\beta_k \mid \text{data})$ under special cases of the prior? Can they be efficiently computed without pSUN sampling?
- Computing: general approach or algorithms tailored to important special cases? software?
- Extensions?
  - Semiparametric model: nonparametric scale normal mixture for the inverse link + pSUN prior for $\boldsymbol{\beta}$.
  - General model settings with binary regression as a component (binary longitudinal responses, nonparametric mixtures, ordinal regression ...)
- Many thanks to Brunero for his contribution to the conference, and congrats to Paolo & Brunero for a nice paper!

A few comments

- Expressions for $E(\boldsymbol{\beta} \mid \text{data})$ or $E(\beta_k \mid \text{data})$ under special cases of the prior? Can they be efficiently computed without pSUN sampling?

- Computing: general approach or algorithms tailored to important special cases? software?

- Extensions?

  - Semiparametric model: nonparametric scale normal mixture for the inverse link + pSUN prior for $\boldsymbol{\beta}$.

  - General model settings with binary regression as a component (binary longitudinal responses, nonparametric mixtures, ordinal regression ...)

- Many thanks to Brunero for his contribution to the conference, and congrats to Paolo & Brunero for a nice paper!

A few comments

- Expressions for $\mathrm{E}(\boldsymbol{\beta} \mid \text{data})$ or $\mathrm{E}(\beta_k \mid \text{data})$ under special cases of the prior? Can they be efficiently computed without pSUN sampling?

- Computing: general approach or algorithms tailored to important special cases? software?

- Extensions?
  - Semiparametric model: nonparametric scale normal mixture for the inverse link + pSUN prior for $\boldsymbol{\beta}$.
  - General model settings with binary regression as a component (binary longitudinal responses, nonparametric mixtures, ordinal regression ...)

- Many thanks to Brunero for his contribution to the conference, and congrats to Paolo & Brunero for a nice paper!